This document is from https://mathematicaster.org/teaching/advgraphs2025/hw/ws4.pdf

Problem 1 (Median vs. Mean)**.** Let X be a random variable.

A. Prove that if X is non-negative, then $\mathbb{M}\,X \leq 2\,\mathbb{E}\,X$.

B. Prove that $|\mathbb{M}\,X - \mathbb{E}\,X| \leq \sqrt{2\,\mathbf{Var}\,X}$.

Now suppose that we have decent tail-bounds centered at some number m in the sense that

$$\int_0^\infty \mathbf{Pr}\big[|X-m| \geq t\big]dt \leq K$$

where K is finite.

C. Prove that $|\mathbb{E}\,X - m| \leq K$.

D. Prove that $|\mathbb{M}\,X - m| \leq 2K$.

E. What does this tell us about the relationship between the median and mean of X in the various concentration inequalities we have proved? (e.g. McDiarmid, Talagrand)

Problem 2 (Max Spanning Tree)**.** Let G be a connected graph on n vertices. For each edge of G, independently sample some weight in $[0,1]$ (possibly different distributions per edge) and let T be the weight of the heaviest spanning tree. Prove that

$$\mathbf{Pr}\big[|T - \mathbb{M}\,T| \geq t\sqrt{n}\big] \leq Ce^{-ct^2}.$$

Compare this to the concentration result given by McDiarmid.

N.b. It turns out that this is a pretty bad bound in general. In the case of $G = K_n$ with weights sampled uniformly from $[0,1]$, we have $\mathbb{E}\,T = n - 1 - \zeta(3) + o(1)$ and $\mathbf{Pr}\big[|T - \mathbb{E}\,T| \geq \epsilon\big] = o(1)$.[1]

Problem 3 (Last Passage Percolation)**.** Let G be a directed graph and fix vertices u, v such that the distance from u to v is at most ℓ (but could be much shorter). For each edge of G, independently sample some weight in $[0,1]$ (possibly different distributions per edge) and let P be the weight of the heaviest u-v path using at most ℓ edges. Prove that

$$\mathbf{Pr}\big[|P - \mathbb{M}\,P| \geq t\sqrt{\ell}\big] \leq Ce^{-ct^2}.$$

Show that McDiarmid gives a similar result under the additional assumption that G is acyclic. **Oh no! I lied! You actually need some sort of "grading" condition on G in order to use McDiarmid...**

Problem 4 (Concentration of Eigenvalues)**.** Let A be a random symmetric matrix with entries in $[-1,1]$ where the entries in the upper-triangle are independent (with any distributions). Let $\lambda_k(A)$ denote the kth largest eigenvalue of A. Prove that

$$\mathbf{Pr}\big[|\lambda_k(A) - \mathbb{M}\,\lambda_k(A)| \geq tk\big] \leq Ce^{-ct^2}$$

Hint #1: (Standard linear algebra exercise) If $A \in \mathbb{R}^{n\times n}$ is symmetric with orthogonal eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$, then

$$\lambda_k(A) = \max_{\substack{\vec{x}\perp \mathrm{span}\{\vec{v}_1,\ldots,\vec{v}_{k-1}\}: \\ \|\vec{x}\|=1}} \vec{x}^T A\vec{x} \qquad \text{and} \qquad \lambda_k(A) = \min_{\substack{\vec{x}\in \mathrm{span}\{\vec{v}_1,\ldots,\vec{v}_{k}\}: \\ \|\vec{x}\|=1}} \vec{x}^T A\vec{x}.$$

Hint #2: Let $\vec{v}_1, \ldots, \vec{v}_n$ be orthogonal eigenvectors for A. For another matrix B, consider unit vectors in $\mathrm{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$ which are orthogonal to the first $k-1$ eigenvectors of B.

[1] https://doi.org/10.1016/0166-218X(85)90058-7